5 SIGN TEST: CONTAMINANT NOT PRESENT IN BACKGROUND

The statistical test discussed in this section is used to compare each survey unit directly with the applicable release criterion. With only the set of survey unit measurements being analyzed, the Sign test used here is called a one-sample test. This section applies if

(1) radionuclide-specific measurements are made to determine the concentrations, and

(2) the background concentration of the radionuclide is negligible.

Otherwise, the methods of Chapter 6 and 7 should be used. Together the above conditions eliminate the need for a reference. The residual radioactivity concentrations in the survey unit are compared directly to the $DCGL_W$ value. The background concentration of the radionuclide need not be zero, but this background amount will be included with the residual radioactivity when analyzing the survey results. The amount that is considered negligible depends on the fraction of the $DCGL_W$ that it represents, and how much residual radioactivity is actually in the survey unit. The risk of the survey unit failing because the background concentration of the radionuclide is included with the residual radioactivity total should be weighed against the savings obtained by not having to make reference area measurements. Sites need not be contiguous areas, however the statistical test are generally applied to individual survey units that cover contiguous areas.

The Sign test is designed to detect uniform failure of remedial action throughout the survey unit. This test does not assume that the data follow any particular distribution, such as normal or log-normal. In addition to the Sign Test, the $DCGL_{EMC}$ for the Elevated Measurement Comparison (EMC)—described in Chapter 8—is compared to each measurement to ensure none exceeds the $DCGL_{EMC}$. If a measurement exceeds this $DCGL_{EMC}$, then additional investigation is recommended—at least locally—to determine the actual areal extent of the elevated concentration.

5.1 Introduction

The use of the Sign test in Scenario A and Scenario B is described in the next two sections, illustrated with example data. The same data will be used in both cases. We consider a survey unit that has been remediated, but may have some residual radioactivity. The $DCGL_W$ for the radionuclide in question has been determined to be 15.9. (The particular radionuclide and units of measurement are irrelevant to the example, and will be left arbitrary.) From data collected during the remediation, it is estimated that the standard deviation of the measurements in the survey unit made during the final status survey should be about 3.3.

During the DQO process it was agreed that the decision error rates should be both set equal to 0.05 initially, and determine if a reasonable survey design could meet these. It was estimated that the costs of additional remediation would be moderate down to concentrations of about 11.5, but would rise sharply below that. On the basis of these considerations, a chart of the desired probability that the survey unit passes was developed and is shown in Figure 5.1.

The lower bound of the gray region is 11.5, and the $DCGL_W$ is 15.9, so $\Delta = 15.9 - 11.5 = 4.4$. Since σ is estimated at 3.3, $\Delta/\sigma = 4.4/3.3 = 1.3$. From Table 3.2, for $\alpha = \beta = 0.05$, 21 samples are required.

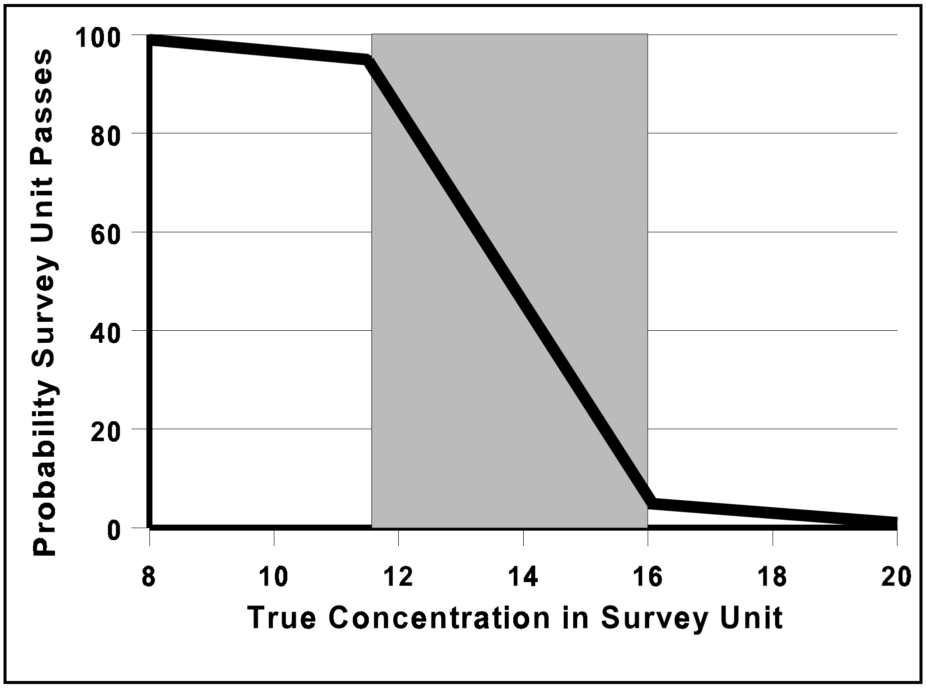

Figure 5.1 Desired Probability That the Survey Unit Passes

The data were taken on a triangular grid. The posting plot is shown in Figure 5.2. It is clear from this plot that there is residual radioactivity that is higher near the center of the survey unit and that diminishes as the survey unit boundary is approached. It is near zero on the west side.

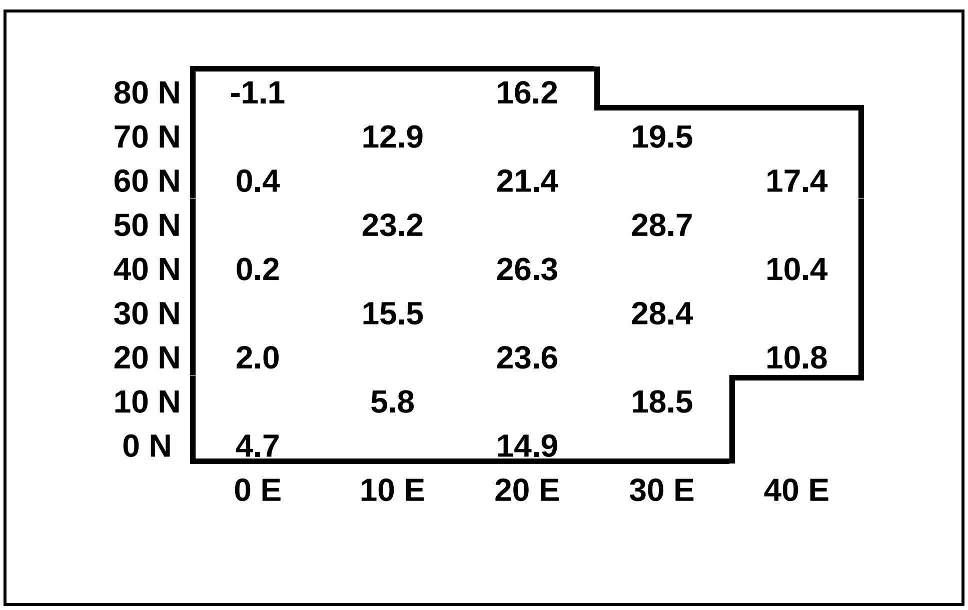

Figure 5.2 Posting Plot of Survey Unit Data

The one reported negative value stands out. Negative values can occur whenever an instrument, analysis or blank background is subtracted to obtain a net reading for a sample[1], which is then converted to a concentration.

Summary statistics for these data are shown in Table 5.1.

Table 5.1 Summary Statistics for Example Data of Figure 5.2

Mean	14.3
Standard Error	2.07
Median	15.5
Standard Deviation	9.5
Sample Variance	90.0
Kurtosis	−1.09
Skewness	−0.18
Range	29.8
Minimum	−1.1
Maximum	28.7
Count	21

The mean is 14.3, which in this case is actually less than the median, 15.5. This is the influence of the few low values on the west part of the survey unit, and is also the cause for the slight negative skewness. The range, 29.8, is only about three times the standard deviation, 9.5. This standard deviation, however, is three times larger than the value of 3.3 assumed in the survey design. If the null hypothesis is accepted, the effect of this higher standard deviation on the power of the test should be investigated. Recall that lower power in Scenario A means that more survey units with concentrations that are actually lower than the $DCGL_W$ are apt to fail the test. In Scenario B, lower power means that more survey units with concentrations greater than the $DCGL_W$ are apt to pass the test. A retrospective power curve, calculated as described in Chapter 10, can be used to decide if the error rates achieved are acceptable. A retrospective power analysis is not necessary when the null hypothesis is rejected, since the Type I error rate, α, is fixed at the design value when the critical value for the test is determined. However, taking additional samples to increase the power will increase the Type I error rate, unless provision for a two-stage test is made during the survey design. This is discussed further in Section 14.2.

A histogram of the data is shown in Figure 5.3. Except for the one negative value in the “zero” bin, the distribution appears reasonably symmetric.

---

(1) For samples with no residual radioactivity, negative values should occur about 50% of the time. It does not imply the existence of negative concentrations. Due to random fluctuations, the measurement merely happened to be lower during the sample analysis than it was during the background determination.

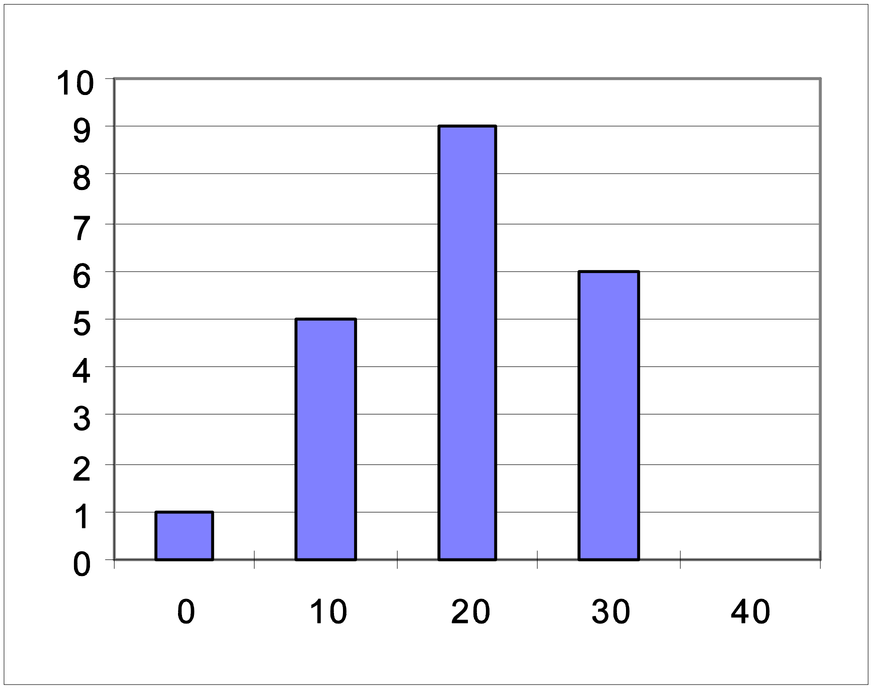

Figure 5.3 Histogram of Survey Unit Data

Further insight into the distribution of the data can be obtained using a ranked data plot. Table 5.2 shows the survey unit data ranked in increasing order. From this table it is already apparent that there is a large gap between the 6th and 7th ranked values. This is even clearer in the ranked data plot of Figure 5.4. It seems that there is a mixture of two distributions in the survey unit data. One is a distribution of low values near the western edge of the survey unit, and one of higher values in the rest of the survey unit. This partially explains the negative skewness and the large standard deviation.

Table 5.2 Ranked Data for Example of Figure 5.2

Rank	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Measurement	−1.1	0.2	0.4	2	4.7	5.8	10.4
Rank	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Measurement	10.8	12.9	14.9	15.5	16.2	17.4	18.5
Rank	**15**	**16**	**17**	**18**	**19**	**20**	**21**
Measurement	19.5	21.4	23.2	23.6	26.3	28.4	28.7

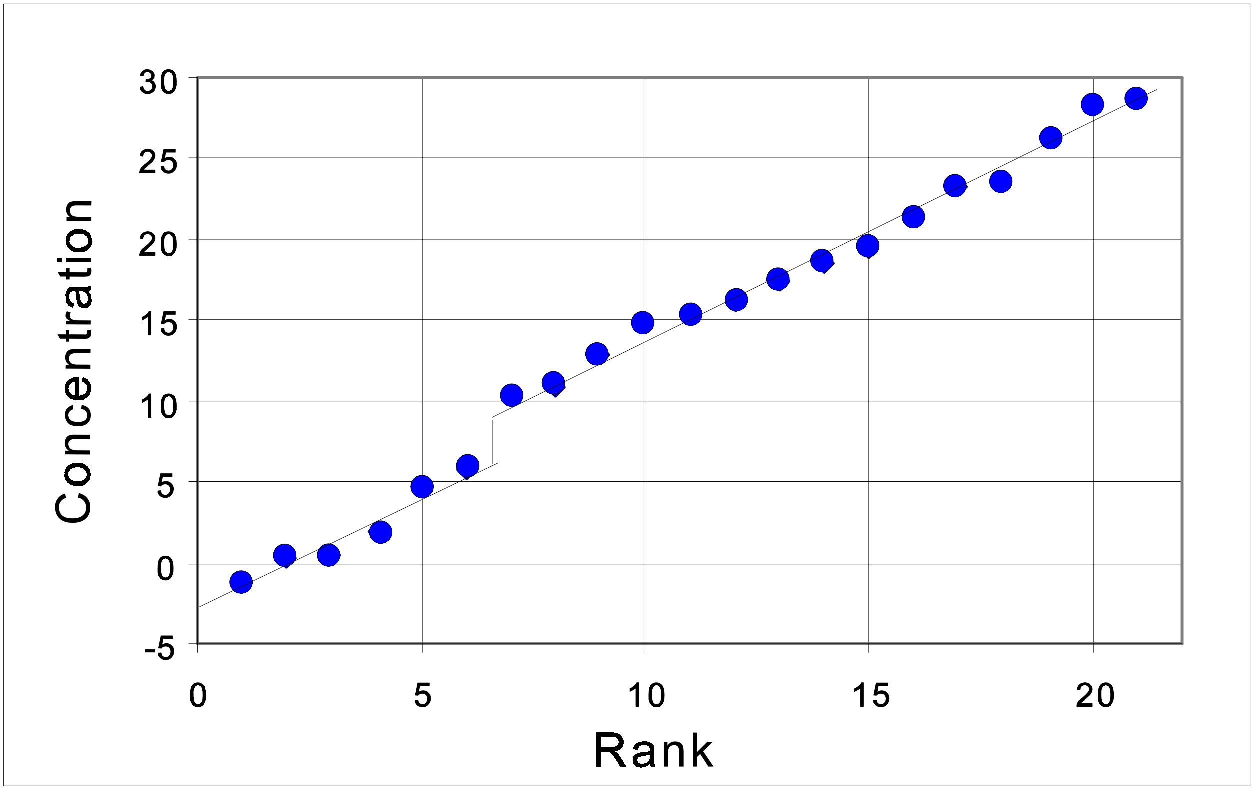

Figure 5.4 Ranked Data Plot of Survey Unit Data

5.2 Applying the Sign Test: Scenario A

The hypothesis tested by the Sign test under Scenario A is:

Null Hypothesis:

H_0: The median concentration of residual radioactivity in the survey unit is greater than the $DCGL_W$.

versus

Alternative Hypothesis:

H_a: The median concentration of residual radioactivity in the survey unit is less than the LBGR.

In order to use the one-sample Sign test, background concentrations of the radionuclide of concern are considered to be either zero or insignificant in comparison to the $DCGL_W$. Thus, there is no reference to background in statement of the null and alternative hypothesis. The null hypothesis is assumed to be true unless the statistical test indicates that it should be rejected in favor of the alternative. The parameter of interest is the mean concentration. The median is equal to the mean when the measurement distribution is symmetric, and is an approximation otherwise.

SIGN TEST

The null hypothesis states that the probability of a measurement less than the $DCGL_W$ is less than one-half, i.e., the 50th percentile (or median) is greater than the $DCGL_W$. The median is the concentration that would be exceeded by 50% of the measurements. Note that some individual survey unit measurements may exceed the $DCGL_W$ even when the survey unit as a whole meets the release criterion. In fact, a survey unit that averages close to the $DCGL_W$ might have almost half of its individual measurements greater than the $DCGL_W$. Such a survey unit may still meet the release criterion.

The hypothesis specifies a release criterion in terms of a $DCGL_W$ which is calculated as described in Section 3.3. The test should have sufficient power ($1-\beta$, as specified in the DQOs) to detect residual radioactivity concentrations at the Lower Boundary of the Gray Region (LBGR). If σ is the standard deviation of the measurements in the survey unit, then Δ/σ expresses the size of the shift (i.e., $\Delta = DCGL_W - LBGR$) as the number of standard deviations that would be considered large for the distribution of measurements in the survey unit. The procedure for determining Δ/σ was given in Section 3.8.1.

The Sign test is applied as follows in Scenario A:

(1) List the survey unit measurements, X_i , i = 1, 2, 3..., N. If a measurement is listed as “less than” a given value, insert that value for the measurement.

(2) Subtract each measurement, X_i , from the $DCGL_W$ to obtain the differences: $D_i = DCGL_W - X_i$, i = 1, 2, 3..., N.

(3) If any difference is exactly zero, discard it from the analysis, and reduce the sample size, N, by the number of such zero measurements.

(4) Count the number of positive differences. The result is the test statistic S+. Note that a positive difference corresponds to a measurement below the $DCGL_W$ and contributes evidence that the survey unit meets the release criterion.

(5) Large values of S+ indicate that the null hypothesis is false. The value of S+ is compared to the table of critical values in Section A.3. If S+ is greater than the critical value, k, in that table, the null hypothesis is rejected.

For the example survey unit data, the calculations are shown in Table 5.3. Notice that when the data are ranked, it is really only necessary to observe the rank of the smallest measurement below the $DCGL_W$ in order to determine S+. However, the differences $D_i = DCGL_W - X_i = 15.9 - X_i$ are also shown. The number of positive differences, S+ = 11. The critical value of S+ for the Sign Test with $\alpha = 0.05$ and $N = 21$ is 14. Since S+ is less than 14, the null hypothesis cannot be rejected. The survey unit has failed the test.

Table 5.3 Calculations for Sign Test in Scenario A

Rank	Measurement	$DCGL_W$ − Measurement	Sign
1	−1.1	14.8	+
2	0.2	15.7	+
3	0.4	15.5	+
4	2.0	13.9	+
5	4.7	11.2	+
6	5.8	10.1	+
7	10.4	5.5	+
8	10.8	5.1	+
9	12.9	3.0	+
10	14.9	1.0	+
11	15.5	0.4	+
12	16.2	−0.3	−
13	17.4	−1.5	−
14	18.5	−2.6	−
15	19.5	−3.6	−
16	21.4	−5.5	−
17	23.2	−7.3	−
18	23.6	−7.7	−
19	26.3	−10.4	−
20	28.4	−12.5	−
21	28.7	−12.8	−

5.3 Applying the Sign Test: Scenario B

The hypothesis tested by the Sign test under Scenario B is:

Null Hypothesis:

H_0: The median concentration of residual radioactivity in the survey unit is less than the LBGR.

versus

Alternative Hypothesis:

H_a:The median concentration of residual radioactivity in the survey unit is greater than the $DCGL_W$.

In order to use the one-sample Sign test, background concentrations of the radionuclide of concern are considered to be either zero or insignificant in comparison to the LBGR. Thus, there is no reference to background in statement of the null and alternative hypothesis. The null hypothesis is assumed to be true unless the statistical test indicates that it should be rejected in favor of the alternative. The Type I error rate, α, is the probability that a survey unit with residual radioactivity at the LBGR will fail to be released. The power, $1 - \beta$, is the probability that a survey unit with residual radioactivity at the $DCGL_W$ will fail to be released. The parameter of interest is the mean concentration. The median is equal to the mean when the measurement distribution is symmetric, and is an approximation otherwise.

The Sign test is carried out for Scenario B in a manner very similar to that for Scenario A:

(1) List the survey unit measurements, X_i , i = 1, 2, 3..., N. If a measurement is listed as "less than" a given value, insert that value for the measurement.

(2) Subtract the LBGR from each measurement, X_i , to obtain the differences: $D_i = X_i -$ LBGR, i = 1, 2, 3..., N.

(3) If any difference is exactly zero, discard it from the analysis, and reduce the sample size, N, by the number of such zero measurements.

(4) Count the number of positive differences. The result is the test statistic S+. A positive difference corresponds to a measurement above the LBGR and is evidence that the median concentration of residual radioactivity survey unit may exceed it.

(5) Large values of S+ indicate that the null hypothesis is false. The value of S+ is compared to the table of critical values in Section A.3. If S+ is greater than the critical value, k, in that table, the null hypothesis is rejected.

For the example survey unit data, the calculations are shown in Table 5.4. Notice that when the data are ranked, it is really only necessary to observe the rank of the smallest measurement below the LBGR in order to determine S− = N − S+. However, the differences $D_i = X_i -$ LBGR = $X_i - 11.5$ are also shown. The number of positive differences, S+ = 13. The critical value of S+ for the Sign Test with $\alpha = 0.05$ and $N = 21$ is 14. Since S+ is less than 14, the null hypothesis cannot be rejected. The survey unit has passed the test. However, it remains to determine the power of the test. Since the observed standard deviation is much greater than that estimated for the test design, it is likely that this survey unit passed simply because there was insufficient power to detect residual radioactivity at the $DCGL_W$. The power calculation for this example is given in Section 10.1.

5.4 Interpretation of Test Results

Once the results of the statistical tests are obtained, the specific steps required to achieve site release will depend on the procedures described in the regulatory guide. The following are suggested considerations for the interpretation of the test results with respect to the release limit established for the site.

Table 5.4 Calculations for Sign Test in Scenario B

Rank	Measurement	Measurement - LBGR	Sign
1	−1.1	−12.6	−
2	0.2	−11.3	−
3	0.4	−11.1	−
4	2.0	−9.5	−
5	4.7	−6.8	−
6	5.8	−5.7	−
7	10.4	−1.1	−
8	10.8	−0.7	−
9	12.9	1.4	+
10	14.9	3.4	+
11	15.5	4.0	+
12	16.2	4.7	+
13	17.4	5.9	+
14	18.5	7.0	+
15	19.5	8.0	+
16	21.4	9.9	+
17	23.2	11.7	+
18	23.6	12.1	+
19	26.3	14.8	+
20	28.4	16.9	+
21	28.7	17.2	+

5.4.1 If the Null Hypothesis Is Not Rejected

Whenever the null hypothesis is not rejected, it is important to complete the analysis by performing a retrospective power analysis for the test. In Scenario A, this will ensure that further remediation is not required simply because the final status survey was not sensitive enough to detect residual radioactivity below the LBGR. In Scenario B, this will ensure that a survey unit is not released simply because the final status survey was not sensitive enough to detect residual radioactivity above the guideline level. The power analysis may be performed as indicated in Chapter 10, using the actual values of the number of measurements, N, and their observed measurement standard deviation s in place of σ. In some cases, a site specific simulation of the retrospective power may be warranted when sufficient power cannot be demonstrated by any of the other suggested methods.

5.4.2 If the Null Hypothesis Is Rejected

If the null hypothesis for the Sign test is rejected in Scenario A, it indicates that the residual radioactivity in the survey unit is less than the $DCGL_W$. However, it may still be necessary to document the concentration of residual radioactivity. It is generally best to use the average measured concentration for this purpose.

If the null hypothesis is rejected in Scenario B it indicates that the residual radioactivity in the survey unit exceeds the LBGR. In this case it is important to determine not only the average concentration of residual radioactivity in the survey unit, δ, but also whether this amount exceeds the release criteria. When the data are normally distributed, the average concentration is generally the best estimator for δ. However, when the data are not normally distributed, other estimators are often better for the same reasons that nonparametric tests are often better than the corresponding parametric tests. These methods are discussed by Lehmann and D'Abrera (1975) . When the estimate for δ is below $DCGL_W$, the survey unit may be judged sufficiently remediated, subject to ALARA considerations. Otherwise, further remediation will generally be required.

The treatment of data that are below the limit of detection will be an important component of these calculations. Whenever possible, the actual results of the measurement should be reported, with an associated total uncertainty that includes both random and systematic errors. Replacing values below the MDC with the MDC value will cause δ to be overestimated.